Resources, Conservation and Recycling 36 (2002) 175–185

Resources Conservation & Recycling

www.elsevier.com/locate/resconrec

National and regional implications of internet data center growth in the US

Jennifer Mitchell-Jackson [a], Jonathan G. Koomey [b,*], Michele Blazek [c], Bruce Nordman [b]

[a] Opinion Dynamics Corporation, 1030 Massachusetts Ave., Cambridge, MA, 02138, USA
[b] Lawrence Berkeley National Laboratory, University of California, 1 Cyclotron Road, Mail Stop 90-4000, Berkeley, CA 94720, USA
[c] AT&T, Suite 3188, 4430 Rosewood Drive, Pleasanton, CA 94588, USA

Accepted 24 June 2002

Abstract

The electricity consumption of data center hosting facilities (also known as server farms or server hotels) is a growing concern to utility demand forecasters, data center facility managers, energy analysts and policy makers. Combining estimates of US computer room floor space for hosting facilities with total computer room power density, we found that these US data centers in the aggregate required less than 500 MW of power in 2000, and used only about 0.12% of the electricity consumed nationwide in that year. In this paper, our order-of-magnitude estimate suggests that energy demands of these facilities do not represent an enormous new burden on the electricity industry as a whole. The fact that these facilities tend to be concentrated in certain areas, however, may mean that there will be significant regional electricity demands in some parts of the country. If combined heat and power (CHP) technologies were introduced to data center facilities on a large scale, initial calculations indicate that these facilities might even become net contributors of power to the electric grid. Published by Elsevier Science B.V.

Keywords: Data centers; Power; Electricity consumption; Power densities

* Corresponding author. Tel.: +1-510-486-5974; fax: +1-510-486-4247
E-mail address: jgkoomey@lbl.gov (J.G. Koomey).

0921-3449/02/$ - see front matter. Published by Elsevier Science B.V.
PII: S0921-3449(02)00080-0

1. Introduction

Concern has grown recently over the alleged demand growth associated with computers and office equipment (Koomey et al., 2000; Kawamoto et al., 2002; Roth et al., 2002) and that associated with new Internet hosting facilities in particular (Mitchell-Jackson et al., 2001; Stein, 2002). The accuracy of these estimates of future demand has important implications for investments in electrical generation, transmission, and distribution capacity. This paper presents our latest assessment of the electricity demands of US data center hosting facilities, so that expectations of future demand growth in such facilities can be based on measured data and not on the speculation currently commonplace on this topic.

2. National impacts

To estimate total power used by US hosting-type data centers, we rely on two parameters: total floor area associated with computer rooms in these facilities, and the total power used per unit of floor area within those computer rooms. The product of floor area and power per unit floor area is total power in Watts. Since the load shape of these facilities is roughly flat, the total energy consumption is simply the product of total power and the number of hours in 1 year (8760 h).

3. Estimating computer room floor space in US data centers

Only a few rough estimates of total US data center floor space exist, and even fewer are truly indicative of Internet growth. Many estimates of total building floor area for data center facilities include significant amounts of space that is not critical to a data center's main function. In a data center facility, 15–75% of the building's area is usually designated for meeting rooms, offices, restrooms and hallways, but only the computer room floor area is relevant to assessments of the growth in electricity demand of data centers. Computer room floor area includes both the area under the racks of computer equipment and that of the electrically inactive areas, such as the aisles between racks of computer equipment. It does not include office space, lobbies, bathrooms, space set aside for potential future expansion, or mechanical equipment rooms.

There are at least two estimates of computer room floor area for hosting facilities, both of which are based on surveys of large data center hosting companies. One estimate by Salomon Smith Barney anticipated that there would be nearly 900 000 m^2 of data center computer rooms built or under construction in the United States at the end of 2000 (Mahedy et al., 2000). A second estimate by the Yankee Group predicted that there would be 860 000 m^2 of computer rooms at the end of 2000—approximately 6% less than in the Salomon Smith Barney report (Yankee Group, 2000). In this report we use 900 000 m^2 as a reasonable one-significant-figure estimate of floor area for these US facilities at the end of 2000.

Although these are currently the best available estimates, they are still approximate. Neither of the estimates mentioned above includes all of the companies that own hosting-type data centers, thus the aggregate area may be underestimated due to the exclusion of some hosting facilities. In addition, in-house corporate data centers are excluded from these estimates (they are not considered to be hosting space, and they were not driving the concern over data center power use in the late 1990s, while proposed and newly-built hosting facilities were).

It is also possible that these estimates are overstated because some of the data centers under construction were never completed, because of double counting, or because of other definitional or accounting oddities. The Salomon Smith Barney report, for example, indicates that at the end of 2000, HostPro, Inc. had five data centers with approximately 5950 m^2 of computer room space, as well as 8730 m^2 of data center computer room space in 'unspecified' locations (Mahedy et al., 2000). According to the HostPro website, however, the company had only five data centers as of April 2000 (HostPro website, 2001). The HostPro estimate included in the Salomon Smith Barney report, therefore, might be more than twice as large as HostPro's actual computer room floor area.

For these reasons, readers should treat these estimates of current electricity used by data centers as an initial 'rough cut' that will be refined with further data collection in the future. Surprisingly, there is even substantial uncertainty about the year 2000 floor area estimates. Recent economic events in this industry have led to a significant slowdown of data center construction, so the extent of future growth in floor area is also uncertain. More detailed surveys will be needed in the future to resolve these uncertainties.

4. Estimating average power density in data center computer rooms

In detailed work on the energy consumption of data centers, we found that there is much confusion over the power requirements at data center facilities (Mitchell-Jackson et al., 2001; Mitchell-Jackson, 2001). We identified many reasons for the overstated energy requests such as erroneously extrapolating total building power requirements from the power density of an isolated area or using design rather than actual power requirements (Mitchell-Jackson, 2001). In addition, many estimates neglect to include the power necessary for end uses such as heating, ventilation, and cooling (HVAC) or other equipment that is necessary to the functioning of the data center.

Due to the confusion over this issue, in our earlier work we developed a common metric, total computer room power density, which can be used to estimate the power requirements of this industry (Mitchell-Jackson et al., 2001). For the purposes of this article, we define total computer room power density as the power drawn by the computer equipment and all of the supporting equipment such as power distribution units, uninterruptible power supplies, HVAC and lights (in W) divided by the computer room floor area (in square meters or m^2). Total computer room power density is a meaningful indicator of power needs and can be compared between buildings of different sizes as well as between data centers at different stages of development.

To determine an upper limit value for total computer room power density, we reviewed billing data for five data center facilities throughout the country (see Table 1). Electricity billing data were used to find average demand in the month with the highest consumption (this month was usually the most recent one). The highest average power demand for the facility was then divided by the computer room floor area. This estimate of total computer room power density is an overestimate because it assumes that all of the power for the entire facility is used for the computer room; thus our estimate is an upper limit for the total computer room power density. Even these overestimates, however, indicate that the total computer room power density was always less than 430 W/m^2 (40 W/ft^2). For comparison, typical office buildings have peak electrical demands of 54–86 W/m^2 (5–8 W/ft^2). For purposes of the national impact calculations below, we adopt a high-end estimate of total computer room power density of 540 W/m^2 (50 W/ft^2).

Computer room power density is partly a function of the occupancy level of the space and of the racks within the space. The average rack in the Bay area data center that we studied most closely was only one-third full (Mitchell-Jackson et al., 2001). Unfortunately, we do not have similar occupancy data for the other four data centers analyzed here, but our experiences in visiting other data centers and in talking with people who design such facilities leads us to conclude that the occupancy level for that Bay area data center was probably typical for hosting-type facilities.

Assuming that approximately 900 000 m^2 of hosting-type data center space was devoted to computer equipment in the US in 2000, and using a high-end average estimate for total computer room power density (540 W/m^2, or 50 W/ft^2), these facilities would require less than 500 MW of power and would use only 0.12% of all electricity used nationwide (see Table 2). Moreover, it is important to note that some unknown portion of this demand is not actually new electricity demand. Some of the computers in these data centers have recently been relocated from office buildings to data center hosting facilities. In fact, a major market for some smaller data centers is convincing small businesses to move their Internet equipment out of offices and closets and into the safe and secure environment of a hosting facility. In these situations, the previously dispersed computer equipment is often moved a short distance from the office where it was originally located. Thus, this portion of data center electricity use does not represent new electricity demand, merely shifted loads.

5. Regional impacts

While data centers represent only a fraction of the total electricity consumed in the United States, regional electricity demands may be more significant. Despite the decentralized nature of the Internet, the communication system involved with Internet transactions have led to a concentration of electricity demands at locations along the nation's fiber optic backbones. According to a Robertson Stephens report, there were approximately 320 data center hosting facilities in the United States at the

Table 1
Comparison of power densities at five US data centers at the end of 2002

Location	Units	Data center A	Data center B	Data center C	Data center D	Data center E
Building area	m^2	11 643	10 684	14 321	NA	33 292
Computer room area	m^2	2555	3716	4181	4476	3577
Building power density	W/m^2	118	31	106	NA	40
Upper limit for total computer room power density	W/m^2 (W/ft^2)	355 (33)	88 (8)	363 (34)	412 (38)	376 (35)

The upper limit for total computer room power density for data centers B–E was calculated by dividing the average power demand for the entire facility (from billing data) by the computer room area. This number includes all of the power used by the entire building and is, therefore, an overestimate. The estimate for data center A is based on a more detailed study of power requirements (see Mitchell-Jackson et al., 2001).

Table 2
Nationwide electricity demands for data center hosting facilities

	Units	2000
Computer room floor area	Million m^2 (Million ft^2)	0.9 (9.5)
Total computer room power density	W/m^2 (W/ft^2)	538 (50)
Data center total power	MW	475
Data center electricity use	TWh	4.2
US electricity use [a]	TWh	3364
Data centers as % total use	%	0.12%

[a] Total US electricity use from the Energy Information Administration's Annual Energy Outlook 2001 (US DOE, 2000).

end of 2000 (Juarez et al., 2001). This was about 45 more than Salomon Smith Barney's projections (Mahedy et al., 2000) but the Robertson Stephens report (Juarez et al., 2001) covered 60 companies while the Salomon Smith Barney report surveyed only 40. The vast majority of these hosting facilities are clustered in major metropolitan areas such as Atlanta, Austin, Boston, Chicago, Dallas, Denver, Los Angeles, New York City, Phoenix, San Diego, the San Francisco Bay Area, Seattle, and Washington, DC/Northern Virginia.

Silicon Valley, California (consisting of four Bay Area counties: Alameda, San Francisco, San Mateo, and Santa Clara) is one of the largest data center hubs in the country. The report by Robertson Stephens estimates that there are 54 data center hosting facilities in this region (Juarez et al., 2001). According to this report, the hosting facilities in the Bay Area make up approximately 17% of all major hosting facilities in the United States. The Salomon Smith Barney report (which pre-dates the Robertson Stephens report) agrees that approximately 15% of the data centers are housed in this region, but indicates that these data centers may represent less than 15% of the country's data center computer room floor area, thus many Silicon Valley data centers may be smaller than average. This report also indicates that computer room floor space in the New York City area may rival the amount of floor space in the Bay Area (Mahedy et al., 2000; we use 'Bay Area' and 'Silicon Valley' interchangeably, here and below).

In the Bay Area, estimates indicate that there were 93 000 m^2 of data center space at the end of 2000 (Mahedy et al., 2000; this estimate does not include the 200 000 (gross) m^2 US Dataport facility planned for the area). Assuming a power density of 540 W/m^2 (50 W/ft^2), data centers could require 50 MW of power in the Bay Area alone—approximately 10% of the total demanded by data centers nationwide. In the Bay Area, this could mean approximately 440 GWh of electricity a year, or roughly 1.2% of electricity consumption in this area (Mahedy et al., 2000).

In spite of widespread reporting to the contrary, California's energy crisis of 2000 and 2001 was not the result of Internet data center growth. For example, a Computer World article published in January 2001 stated that power demands 'skyrocketed by 12%' in the heart of Silicon Valley compared with a statewide growth rate of 2 or 3%, and implied that this growth was the result of the Internet (Hall, 2001). However, the

Table 3
Electricity consumption in Silicon Valley versus the state of California

Year	Silicon valley total electricity consumption		Statewide total electricity consumption	
	Million kW h	% Growth	Million kWh	% Growth
1990	31 436		228 038	
1991	31 140	−1	222 260	−3
1992	31 587	1	226 988	2
1993	31 679	0	227 624	0
1994	31 467	−1	230 097	1
1995	32 220	2	230 990	0
1996	32 911	2	239 168	4
1997	34 469	5	247 437	3
1998	34 289	−1	244 510	−1
1999	35 360	3	252 800	3
2000	36 616	4	264 429	5
Overall growth 1990–2000		16		16
Average annual growth		1.6		1.5

Source: California Energy Commission (Web site on Silicon Valley Electricity Consumption, 2001). Silicon Valley is defined here to consist of four Bay Area counties: Alameda, San Francisco, San Mateo, and Santa Clara.

data presented on the California Energy Commission's web page on 'Silicon Valley Electricity Consumption' (seen in Table 3) shows that total electricity use in the Silicon Valley did not grow at a substantially higher rate than in the rest of California (Hall, 2001; Web site on Silicon Valley Electricity Consumption, 2001). Both Silicon Valley and statewide electricity consumption grew by roughly 16% over this 11 year period.

While not responsible for California's power crisis, data centers may exacerbate power supply problems in utility service territories where they are concentrated. Data centers appear to be ideal electricity customers: they demand a relatively steady amount of power 24 h a day. In reality, however, even after these facilities are built, utilities do not always have an accurate estimate of demand because this is a new and rapidly changing industry. For a utility, lacking knowledge of the data center's electricity demand leads to difficulties in providing supply and hedging against risks. Utilities also face the challenge of meeting the customer's demand for infrastructure. This problem is especially difficult given the distinct differences in timing and planning cycles between utilities and the data center developers. Utilities are accustomed to getting 2 or 3 years notice for new large office buildings and industrial centers. Now, they face the challenges of putting in power lines, transformers, and substations within a few months.

Faced with new requests for power, utilities must decide how to respond. Building new infrastructure and acquiring new power resources to meet demand will have serious costs. If priced correctly, these costs can be recouped through electricity charges, but if demand is significantly overstated, the utilities will spend more on infrastructure than they could ever hope to recover. Furthermore, utilities may build infrastructure to meet the power demands of an industry that could fade as rapidly as it appeared. In order to avoid excessive risk, several utilities have started to charge data centers for electricity infrastructure based on their initial requests for power. The utilities then pay the developers back in portions over time, depending on how a site's electricity usage progresses towards the original load estimate (Ahlberg, 2001; Cook, 2000).

6. The potential for combined heat and power in data centers

One policy question that has arisen recently is the potential for combined heat and power (CHP) technologies to supply cooling and electricity to data centers. In many ways, these facilities represent ideal loads for CHP, in that they run every hour of the year, with relatively small variations in electric loads and cooling needs. Table 4 shows a simple and illustrative calculation based on the data presented above. The calculation begins with total electricity used by data centers from Table 2, and estimates the contribution of different end-uses to that total, based on the

Table 4
Illustrative total CHP potential to meet data center cooling loads and generate electricity

	Units	2000
Total electricity used by data centers[a]	TWh.e	4.2
Computer room, including lights and aux equipment	TWh.e	2.6
Central chiller plant	TWh.e	0.6
Fans, computer room AC units, other ventilation	TWh.e	1.0
Total cooling load	TWh.th	2.6
Total heat needed to meet cooling load using absorption cooling[b,e]	TWh.th	3.7
Total electricity generated from CHP[c,d,f]	TWh.e	4.9
% of non-chiller electricity generated by CHP		137%
Total capacity (gas industrial IC CHP)	GW.e	0.6

Note: TWh.e, terawatt-hours of electricity; TWh.th, terawatt-hours of thermal energy (heat); GW.e, gigawatts of electricity generation capacity; IC, internal combustion; COP, coefficient of performance; AC, air conditioning.

[a] Total electricity used by data centers taken from Table 2. Breakdown by end use extrapolated from detailed work on electricity consumption in one Bay Area data center from Mitchell-Jackson et al. (2001).

[b] COP of absorption cooling, cooling out/heat in = 0.7. Assumes indirect-fired single-effect absorption cooling.

[c] EPR = electricity out/useful heat out, for a 25 MW industrial IC engine = 1.32. From Krause et al. (1994), TLF, p. A.10.7.7.

[d] Capacity factor (industrial IC CHP) = Equivalent availability = 92%. From Krause et al. (1994), p. A.10.7.7.4

[e] Total heat needed to meet cooling load = Cooling load/COP.

[f] Total electricity generated = total heat needed to meet cooling load × EPR. Assumes a gas fired industrial IC engine, 25 MW electrical capacity, operated in thermal load following mode, taken from Krause et al. (1994).

characteristics of the data center floor area for the Bay Area data center explored in detail in Mitchell-Jackson et al. (2001). We then calculate the cooling load as the amount of heat from computers, lights, and auxiliary equipment, and use a typical coefficient of performance (COP) for absorption cooling to determine the amount of heat needed from the CHP unit to remove the heat represented by the cooling load from the building. Finally, we multiply the heat needed to supply absorption cooling by the electricity production rate (EPR) for a gas-fired internal combustion (IC) engine cogeneration unit to obtain the amount of electricity generated. The EPR characterizes the amount of electricity generated from a CHP plant divided by the amount of useful heat from that same plant (for details on definitions and performance characteristics of typical CHP systems, see Krause et al., 1994).

Table 4 shows that a significant potential for CHP exists for US hosting-type data centers. Total potential CHP generation in 2000 is about 5 TWh/year, which corresponds to about 0.6 GW of capacity. The electricity generated by the CHP is more than enough to supply the non-chiller loads in these data centers, so these facilities have the potential to become net producers of electricity. Of course, local air pollution concerns may influence whether such installations are feasible or desirable. Fuel cells may mitigate some concerns about local air pollution, but more research is needed to determine the cost and operating characteristics of this technology in this application.

7. Conclusions

Based on our findings for total computer room power density and the estimate of computer room floor area for 2000, data center facilities in the aggregate require less than 500 MW of power, and use a fraction of a percent (less than 0.12% in 2000) of all electricity used nationwide. This is less than one-tenth of the electricity consumed by the US iron and steel industry (Margolis and Brindle, 2000). Moreover, some unknown portion of this demand is not actually new electricity demand. Some of the computers in these data centers are just relocated from office buildings to data center hosting facilities. The electricity use by US data centers, therefore, does not represent an enormous new burden on the energy industry, although there may be regional power constraints in areas where data center facilities are concentrated.

Current estimates of vast power needs for data centers need to be revised to reflect the real power demands documented above. Additional detailed studies of current energy requirements, as well as better estimates of total US computer room floor area, will help all parties to better understand the needs of this industry. In particular, more accurate estimates of the amount of floor area attributable to in-house corporate data centers and hosting facilities need to be developed, and real data on the power used by these facilities must be compiled for a significant fraction of the total floor area. Further analysis and data collection are also needed to assess the potential for CHP in these facilities, which our rough calculations above indicate could become net producers of electricity on a large scale. This additional data collection will give us more accurate estimates of total electricity used by and potential generation from data centers, and should be undertaken with all due haste.

Acknowledgements

We would like to acknowledge the support of Skip Laitner of the Environmental Protection Agency, Office of Air and Radiation, and the comments of Ernst Worrell, Skip Laitner, and anonymous journal reviewers. This work was supported by the Office of Air and Radiation of the US Environmental Protection Agency. Prepared for the US Department of Energy under Contract Number DE-AC03-76SF00098.

References

Ahlberg E. Electricity utilities fear drought at server farms. Dow Jones Newswires, 2001 http://www.djnewswires.com.

Cook J. Internet data gain is a major power drain on local utilities. Seattle Post Intelligencer. Seattle, 2000 http://seattlep-i.nwsource.com.

Hall M. Net Blamed as Crisis Roils California. Computer World, 2001 Read on the web at http://www.computerworld.com/cwi/story/0,1199,NAV47_STO56341,00.html.

Information from the HostPro website, 2001. Read on the web on April 12. http://www.hostpro.com.

Juarez RA, Alic MT, Karkhaniz CS, Johnson BD. Space Dex III—Hosting Space: not all space is created equal—smart, complex space takes center stage. Robertson Stephens Inc., Boston, MA, 2001 http://www.rsco.com.

Kawamoto K, Koomey J, Nordman B, Brown RE, Piette M, Ting M, Meier A. Electricity used by office equipment and network equipment in the US energy. Int J (also LBNL-45917) 2002;27:255–69.

Koomey JG. Rebuttal to Testimony on 'Kyoto and the Internet: The energy implications of the digital economy'. Report no. LBNL-46509. Lawrence Berkeley National Laboratory, Berkeley, CA, 2000. http://enduse.lbl.gov/Projects/InfoTech.html.

Krause F, Koomey J, Becht H, Olivier D, Onufrio G, Radanne P. Fossil Generation: the cost and potential of low-carbon resource options in Western Europe. International Project for Sustainable Energy Paths, El Cerrito, CA, 1994.

Mahedy S. Cummins D, Joe D. Internet Data Centers: If Built...Will They Come? Salomon Smith Barney, New York, NY, 2000.

Margolis N, Brindle R. Energy and Environmental Profile of the US Iron and Steel Industry. Report no. DOE/EE-0229. Prepared by Energetics Inc. for the US Department of Energy Office of Industrial Technologies, Washington, DC, 2000.

Mitchell-Jackson J. Energy Needs in an Internet Economy: a Closer Look at Data Centers. Energy and Resources Group. Berkeley, CA: University of California, Berkeley, 2001.

Mitchell-Jackson J, Koomey J, Nordman B, Blazek M. Data Center Power Requirements: measurements from Silicon Valley. Under review by Energy–Int. J. (also LBNL-48554).

Roth K, Goldstein F, Kleinman J. Energy Consumption by Office and Telecommunications Equipment in Commercial Buildings—Volume I: Energy Consumption Baseline. Report no. A.D. Little Reference no. 72895-00. Prepared by Arthur D. Little for the US Department of Energy, Washington, DC. http://www.eren.doe.gov/buildings/documents.

Stein J. More Efficient Technology Will Ease the Way for Future Data Centers. 2002 ACEEE Summer Study on Energy Efficiency in Buildings (Forthcoming), American Council for an Energy Efficient Economy, Asilomar, CA; Washington, DC, 2002.

US DOE. Annual Energy Outlook 2001, with Projections to 2020. Report no. DOE/EIA-0383 (2001). Energy Information Administration, US Department of Energy, Washington, DC, 2000.

Web site on Silicon Valley Electricity Consumption. 2001. Read on the web on March 29. California Energy Commission. http://www.energy.ca.gov/silicon_valley_consumption.html.

Yankee Group, 2000. Executive Summary of The US Collocation Market: High-Tech Real Estate Heats Up. Yankee Group, Boston, MA. http://www.yankeegroup.com.